

## Sistema de puntuación de alerta temprana de descompensación basado en percentiles para pacientes pediátricos hospitalizados

*Centile-based decompensation early warning score system in pediatric inpatients*

Patricio Wolff<sup>1</sup>    Sebastián A. Ríos<sup>1\*</sup>

Recibido 12 de Noviembre de 2019, aceptado 25 de Abril de 2021

*Received: November 25, 2019    Accepted: April 25, 2021*

### RESUMEN

Estudios internacionales muestran que la inclusión de sistemas de alerta temprana (SAT) en los procesos de atención clínica permite a los cuidadores anticipar las descompensaciones de los pacientes y tomar decisiones oportunas. Actualmente, es posible encontrar diferentes perspectivas en la construcción de los SAT, sin embargo, el enfoque basado en percentiles no se ha aplicado a pacientes pediátricos. El objetivo de este estudio es desarrollar un SAT pediátrico utilizando las propiedades estadísticas de los registros históricos de signos vitales. Se desarrolló y evaluó un método de puntuación con un enfoque estadístico supervisado. Se obtuvieron 178,970 registros de signos vitales de 4,104 pacientes de 0 a 18 años, junto con información de registros médicos electrónicos de un hospital pediátrico. Basado en un método de puntuación de 7 bandas y utilizando diferentes signos vitales, se generó un modelo que nos permite predecir una transferencia no planificada de pacientes a la UCI, 8 horas antes de que realmente ocurriera. La predicción obtenida en el mejor de los casos mostró resultados en áreas bajo la curva ROC (AUC) de 0,71. Además, los umbrales de corte calculados son sensibles en las categorías de menor riesgo (0,877 a niveles de bajo riesgo) y poseen una mayor especificidad en las categorías de mayor riesgo (0,89 en la categoría de Alto Riesgo). Los resultados permiten corregir los rangos normales de los signos vitales en el modelo basado en conocimiento experto, que actualmente está en funcionamiento. Nuestro enfoque estadístico supervisado proporciona un método auditable y reproducible.

Palabras clave: SAT, pediatría, Soporte a la decisión clínica.

### ABSTRACT

*International studies show that including Early Warning Score (EWS) Systems in clinical care processes allows caregivers to anticipate patient decompensations and make timely decisions. Currently, it is possible to find different perspectives in the construction of EWS; however, the centile-based approach has not been applied to pediatric patients. This study aims to develop a pediatric EWS using the statistical properties of historical vital sign records. A scoring method with a supervised statistical approach was developed and evaluated. 178.970 vital sign records were obtained from 4,104 patients aged 0 to 18 years, in conjunction with information from Electronic Health Records from a pediatric hospital. Based on a 7-band scoring method and using different vital signs, a model was generated that allows us to predict an unplanned transfer of patients to ICU, 8 hours before it happened. The prediction obtained in the best case showed results in areas under the ROC Curve (AUC) of 0.71. In addition, the calculated cut-off thresholds are sensitive in the lower risk categories (0.877 at low-risk levels) and possessing greater specificity in the higher-risk categories (0.89 in the High-Risk category). The results of this study allow correcting the normal ranges in the vital signs in the Expert-based model, which is currently in operation. Our supervised statistical approach provides an auditable and reproducible method.*

*Keywords: EWS, Pediatric, Clinical Decision Support Systems.*

---

<sup>1</sup> Centro de Investigación en Inteligencia de Negocios, Universidad de Chile, Beauchef 851, Santiago P.O. Box: 8370459, Chile

\* Autor de correspondencia: srios@dii.uchile.cl

## INTRODUCCIÓN

La hospitalización de un paciente involucra distintos riesgos asociados a los tratamientos y cuidados tales como: la ocurrencia de descompensaciones; eventos adversos severos como Paro Cardio Respiratorio (PCR); infecciones intrahospitalarias; caídas; riesgos asociados a la administración de medicamentos; entre otros. La manifestación de estos riesgos tiene consecuencias directas en el proceso de atención, aumentando los tiempos de estadía y los costos de tratamiento. Estudios internacionales muestran que la ocurrencia de eventos adversos severos (SAE) en pacientes hospitalizados puede producir, en promedio, aumentos de 6,1 días en el tiempo de estadía y el doble de costos en relación a pacientes que no sufrieron un SAE [8]. Las descompensaciones, los SAE e infecciones se caracterizan por que una respuesta temprana a ellos minimiza el daño, generando un ahorro de costos y mayor facilidad de los tratamientos. Así mismo, aproximadamente un 41% de los SAE pueden ser evitables con una detección proactiva [11].

Es posible prevenir cierto grupo de descompensaciones en pacientes hospitalizados mediante el control periódico de signos vitales [1]. Estas descompensaciones, en general, están precedidas por signos vitales anormales [9]. Desde 1997 [12], se han desarrollado muchos Sistemas de alerta temprana (SAT) (conocidos como *Early Warning Systems*, EWS) para alertar al personal sobre el posible deterioro de la condición del paciente usando como información central los registros periódicos de signos vitales de los pacientes hospitalizados [4, 7, 9]. El propósito de los SAT es proporcionar una forma accesible y sistemática para evaluar el estado de salud de los pacientes y guiar la respuesta en caso de una situación de deterioro clínico, basado en medidas fisiológicas obtenidas periódicamente durante la hospitalización [2]. Los SAT, en conjunto con el monitoreo de los signos vitales, son las intervenciones de enfermería más utilizada para la detección temprana del deterioro clínico de pacientes hospitalizados [6].

Actualmente, es posible encontrar diferentes perspectivas en la construcción del SAT. Algunos SAT se construyen en base a la opinión de expertos, como el National Early Warning Score (NEWS) [16, 18], el Modified Early Warning Score (MEWS) [19] y el VitalPAC Early Warning Score (VIEWS) [14].

También existen SAT para pacientes pediátricos como el Pediatric Early Warning Score (PEWS) [5], el Bedside PEWS [13] y el Children's Hospital Early Warning Score (C-CHEWS) [10]. Otros modelos de riesgo utilizados en pacientes pediátricos son el Pediatric Index of Mortality (PIM) y el Pediatric Risk of Mortality (PRISM). Estos se centran en la probabilidad de mortalidad del paciente al ingresar a las unidades de cuidados críticos. Estos modelos han presentado desventajas debido al gran número de variables que se requieren, lo que dificulta su utilización en la práctica clínica [15].

En [20] se argumenta que los modelos clínicos basados en conocimiento de expertos pueden mejorarse mediante la estimación estadística de los rangos de normalidad de los signos vitales. El enfoque estadístico no supervisado de [21] permitió mejorar los resultados obtenidos por el centile-based EWS (CEWS) propuesto en [20]. Si bien, esta perspectiva muestra excelentes resultados metodológicamente quedan algunas dudas respecto de la arbitrariedad en la determinación de los percentiles que definen las bandas y los umbrales que determinan el riesgo del puntaje acumulado por paciente.

La validación de estos modelos se puede realizar basándose en la comparación del resultado con la opinión de un equipo de expertos o en base a algún resultado clínico definido. Los SAT mencionados [5, 10, 13, 14, 16, 18, 19] fueron diseñados para detectar pacientes deteriorados en salas de hospitalización, específicamente para prevenir: el ingreso inesperado en la UCI, el regreso no planificado al quirófano, una estadía prolongada, un paro cardíaco o la muerte. En este sentido, la validación de estos modelos debe basarse en la capacidad de predecir estos resultados clínicos [22].

A pesar de que existe una amplia gama de SAT, su rendimiento como muestra [21], depende en gran medida de la base de datos de la que se deriva. Esto ha motivado el desarrollo de diferentes modelos orientados a poblaciones y áreas geográficas específicas.

En la mayoría de los SAT de adultos, el rango de edad de no es una característica central. En caso pediátrico, es relevante conocer el rango de edad del paciente para determinar los niveles de riesgo de la frecuencia cardíaca (FC), la frecuencia respiratoria (FR) y la presión arterial sistólica (PAS), ya que los

signos vitales normales y los valores de referencia fuera de rango de estos grupos son diferentes. Los rangos de edad utilizados se definieron según el soporte vital avanzado pediátrico (PALS) [3].

**Objetivo**

El objetivo de este trabajo es desarrollar un SAT pediátrico, utilizando las propiedades estadísticas de los registros históricos de signos vitales. En este caso particular, los modelos propuestos se evaluarán en base a la capacidad de predecir el traslado de camas básicas a Unidades de Cuidados Intensivos (UCI). Para la predicción se utilizaron los controles de signos vitales periódicos de enfermería.

Las variables seleccionadas para el modelo coinciden con la información disponible en los actuales instrumentos de control de signos vitales, que son llevados a cabo por personal de enfermería y técnicos paramédicos del hospital. Las variables seleccionadas para el modelo son las siguientes: la frecuencia cardíaca (FC), la frecuencia respiratoria (FR) y la presión arterial sistólica (PAS), nivel de conciencia (AVPU), temperatura corporal (Temp.), Saturación de oxígeno (SO) y rango de edad del paciente. Además, se utilizó información sobre la condición del paciente en el momento del registro (Durmiendo, Despierto y Llorando).

Existen otras variables fisiológicas e información del tratamiento que podrían ser incluidas en el modelo como, por ejemplo, requerimientos de oxígeno, recarga capilar (CRT), número de medicamentos administrados [5]. Sin embargo, aumentar el número

de variables registradas complicará la utilización de los sistemas de registro de signos vitales en la práctica clínica. En el caso que esta información exista en otros sistemas (por ejemplo, el Registro Clínico Electrónico, RCE) la operación del modelo se condicionará a la capacidad de integrar ambos sistemas. De cualquier forma, las variables seleccionadas en primera instancia en este trabajo corresponden a la información disponible en el momento del estudio, en la base de datos del sistema de registro de signos vitales. Una observación importante es reportar que, además se disponía de la Presión Arterial Diastólica (PAD) y la Presión Arterial Media (PAM), sin embargo, éstas no mostraron en ninguno de los modelos estudiados una mejora en la capacidad predictiva al incorporarse. Esto último también es reportado en la literatura internacional [20, 21], por lo que estas son las únicas variables que se descartaron del grupo inicial.

**DISEÑO DEL ESTUDIO**

Proponemos un enfoque estadístico supervisado para calcular los percentiles que definen rangos de normalidad en un método de puntuación. Para esto se estudiará la combinatoria de diferentes combinaciones de percentiles por signo vital que determinan niveles de riesgo. Este nivel de riesgo determinado se evaluará en base a la capacidad predictiva de traslado a UCI en menos de 8 horas utilizando las curvas Receiver Operating Characteristic (ROC) y el Área Bajo la Curva ROC (conocido como AUC) en cada grupo de percentiles. Los pasos se describen el esquema de la Figura 1.

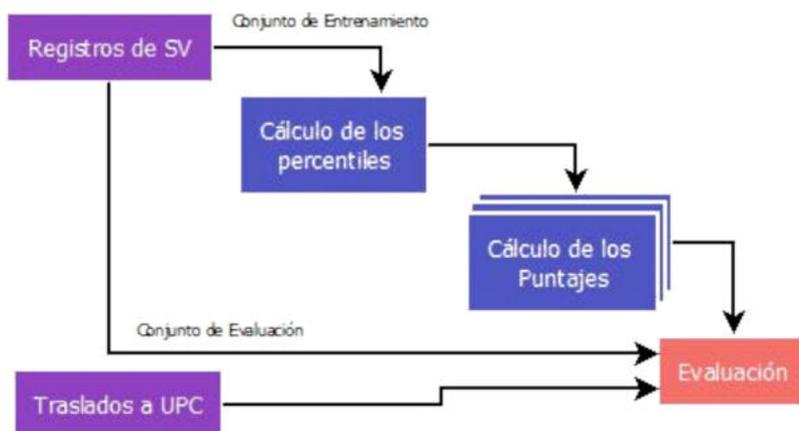


Figura 1. Diseño del Estudio.

### Caracterización del conjunto de datos

En un hospital pediátrico de Santiago de Chile se implementó un sistema móvil que permitió el registro de signos vitales manualmente, en una Unidad de Cuidados Intensivos (UCI), entre el 6 de febrero de 2018 y el 31 de diciembre de 2018. Durante este período, se obtuvieron correctamente 178.970 registros de signos vitales de 4.104 pacientes (2.311 hombres 56,3%) entre 0 y 18 años. La edad media (mediana) de los pacientes fue de 3,2 años. Sobre este conjunto de datos se realizó una partición aleatoria estratificada por edad de pacientes, considerando un 20% de datos para la evaluación y un 80% de entrenamiento. Es decir, los registros de signos vitales se clasificaron según la edad del paciente en el momento del registro. En cada uno de estos grupos se realizó una partición aleatoria de la data en proporciones de 20% para la evaluación y un 80% para el cálculo de los percentiles.

La información de los traslados y su ocurrencia se obtuvo del sistema de gestión de camas del hospital, en igual período. La data que permite evaluar los resultados corresponde a los registros clínicos de pacientes transferidos a la UCI. Del sistema de gestión de camas, se obtuvo un total de 203 (4,95%) pacientes derivados a UCI desde unidades de hospitalización básica. Solo 681 de los registros recopilados se generan 8 horas antes de una transferencia de la UCI (0,38%). Los registros de signos vitales etiquetados como de riesgo corresponden a los tomados 8 horas antes del traslado a UCI.

### Cálculo de percentiles

Los percentiles se pueden estimar a partir de un grupo de  $N$  medidas  $Y_k$  con  $0 < k < N$  y ordenadas de forma que  $Y_k < Y_{k+1}$  para todo  $k$ . Para el percentil  $p$ -ésimo, con  $p \in \{1, 5, 10, 15, 25, 75, 85, 90, 95, 99\}$  se calculan los valores de  $Y_j$  para cada signo vital, tales que:

$$j = \sup(N \cdot p / 100) \quad (1)$$

donde  $\sup(x)$  corresponde al entero inmediatamente superior de  $x$ .

Cada par de  $Y_j$  define un intervalo que a su vez tiene un puntaje asociado  $h$  con  $h \in \{0, 1, 2, 3\}$  o nivel de riesgo de nido en intervalos de siete bandas para cada signo vital. Por definición se asume que

el valor de los extremos de la distribución tendrá el valor  $h$  mayor y 0 en el rango que contienen la mediana de la distribución.

### Cálculo de los puntajes por registro

Para diferentes combinaciones de  $p$  por signo y por nivel de riesgo se evaluó cada signo vital y se sumaron los valores de  $h$  para establecer un puntaje para cada registro de signo vital perteneciente al subgrupo de evaluación. Se probaron 8 combinaciones de  $p$  como se muestra en la Tabla 1.

Es decir, en cada prueba se obtiene un puntaje por cada registro de signos vitales. Luego, la evaluación de la capacidad de predictiva en cada prueba dependerá de la correlación mostrada por cada puntaje calculado con respecto a la etiqueta de riesgo (o registro de signos vitales 8 horas antes de un traslado a UCI).

### Métricas de Evaluación

Este estudio centra la evaluación del modelo de riesgo de descompensación en la capacidad de predecir un traslado no planificado a UCI, dentro de al menos 8 horas. Esta definición se debe a que el reconocimiento de cambios clínicos dentro de 8 horas, junto con una intervención temprana y apropiada, puede prevenir resultados adversos, que incluyen paro cardíaco y muerte en la mayoría de los pacientes (84%) [17]. La transferencia no planificada a la UCI está relacionada con otros resultados (como un paro cardíaco y muerte). Sin embargo, se seleccionó la transferencia no planificada en la UCI como el único resultado, principalmente debido a la baja frecuencia en el hospital de los episodios de muerte y los eventos de paro cardíaco en las salas de hospitalización básica. En este estudio no se utilizó la estancia prolongada, ya que es un resultado menos objetivo (debido a la baja estandarización de la práctica clínica del hospital) que podría dar lugar a errores en los procesos de entrenamiento y la evaluación del modelo.

Para cada prueba y para cada umbral  $u$  con  $u \in \{0, \dots, 18\}$  se calculó la matriz de confusión y las métricas de rendimiento derivadas de ella, que describen a continuación. Si se define  $TP$ ,  $TN$ ,  $FP$  y  $FN$  como verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos respectivamente. Entonces calculamos Sensibilidad ( $Sens$ ) como:

Tabla 1. Detalle de las combinaciones de percentiles utilizados en cada prueba.

|          | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|----------|-------|----------|---------|---------|----------|---------|---------|----------|
| Prueba 1 | FR    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | FC    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | PAS   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 2 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | FC    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 3 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | FC    | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 4 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | FC    | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 5 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | FC    | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 6 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | FC    | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 7 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | FC    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 10   | 10 ~ 15 | 15 ~ 25 | 25 ~ 75  | 75 ~ 85 | 85 ~ 90 | 90 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |
| Prueba 8 | Valor | 3        | 2       | 1       | 0        | 1       | 2       | 3        |
|          | FR    | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | FC    | 0 ~ 5    | 5 ~ 10  | 10 ~ 15 | 15 ~ 85  | 85 ~ 90 | 90 ~ 95 | 95 ~ 100 |
|          | SAT   | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 100 |         |         |          |
|          | Temp  | 0 ~ 1    | 1 ~ 5   | 5 ~ 10  | 10 ~ 90  | 90 ~ 95 | 95 ~ 99 | 99 ~ 100 |
|          | AVPU  | No resp. |         | Palabra | Alerta   |         |         |          |

$$Sens_u = \frac{TP}{TP + FN} \quad (2)$$

la Especificidad como (*Spec*) como:

$$Spec_u = \frac{TN}{TN + FP} \quad (3)$$

la Precisión como (*Prec*) como:

$$Prec_u = \frac{TP}{TP + FP} \quad (4)$$

el *f-score* se define como:

$$f - score_u = \frac{2 \cdot Prec \cdot Spec}{Prec + Spec} \quad (5)$$

El análisis utilizando las curvas *Receiver Operating Characteristic* (ROC) se han utilizado ampliamente para comparar diferentes clasificadores binarios. La ROC es una medida de la sensibilidad frente a la tasa de falsos positivos. Es ampliamente utilizada para comparar los rendimientos del estado del arte en métodos de clasificación. El Área Bajo la Curva ROC (AUC), se reporta a menudo en estudios que compara la capacidad predictiva de diferentes modelos.

## RESULTADOS

Los resultados obtenidos se muestran en la Figura 2. En este gráfico es posible observar que la Prueba 6 entregó mejores resultados con AUC = 0,706 desde el punto de vista de su capacidad predictiva.

En la Figura 3 se muestra la curva ROC obtenida en la prueba 6. Esta curva se obtiene al determinar la sensibilidad y 1-especificidad para cada umbral entre 0 y 18.

En general, los SAT poseen 4 niveles de riesgo como: Sin Riesgo, Riesgo Bajo, Riesgo Medio y Riesgo Alto. Para determinar cuál es el riesgo para cada puntaje se requiere determinar umbrales o puntajes de corte. Para este análisis se propone evaluar (para los puntajes obtenidos en la Prueba 6) el comportamiento de las métricas descritas en la sección 2,4 recorriendo todos los umbrales posibles (entre 0 y 18).

Gráficamente, es posible observar en la Figura 4 el comportamiento de las métricas de evaluación en función de un determinado puntaje de corte. En base a este comportamiento se determinarán 3 umbrales que definen los puntajes en cada nivel de riesgo (ver detalle en el Apéndice A).

Desde el punto de vista clínico, se prefiere un método de predicción de alta sensibilidad. Sin embargo,

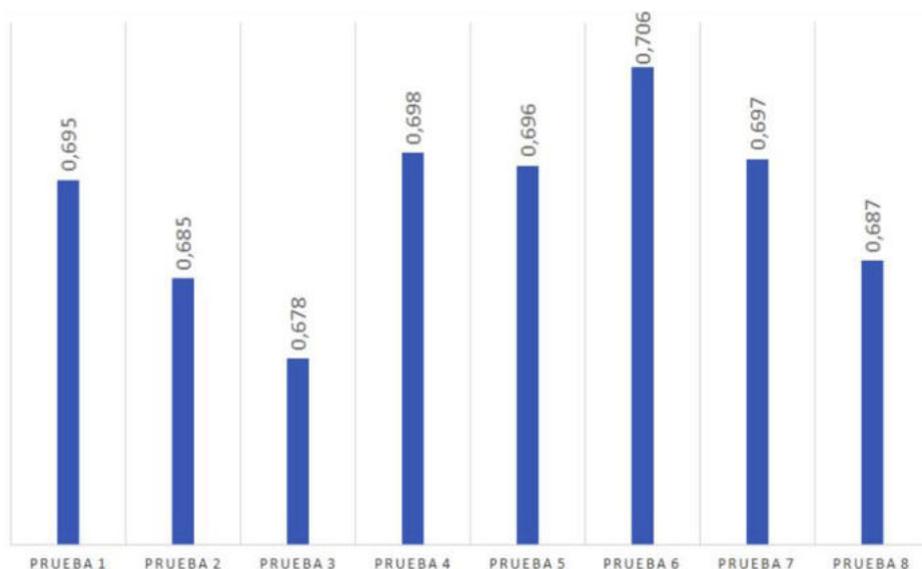


Figura 2. AUC obtenidos.

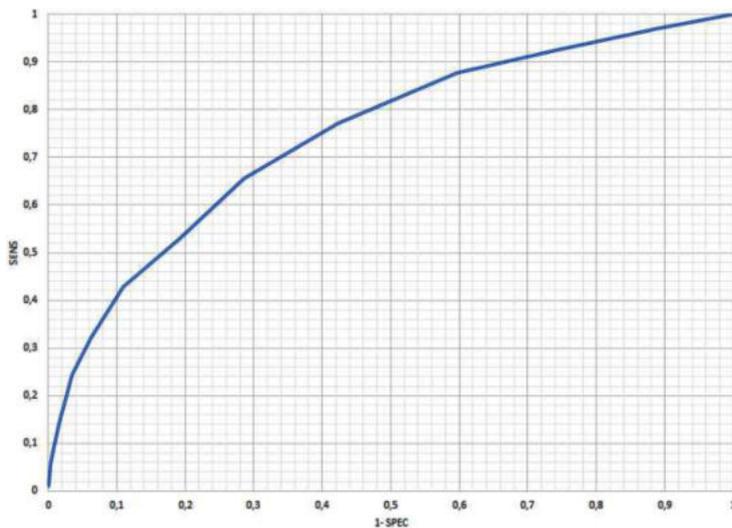


Figure 3. Curva ROC de la Prueba 6.

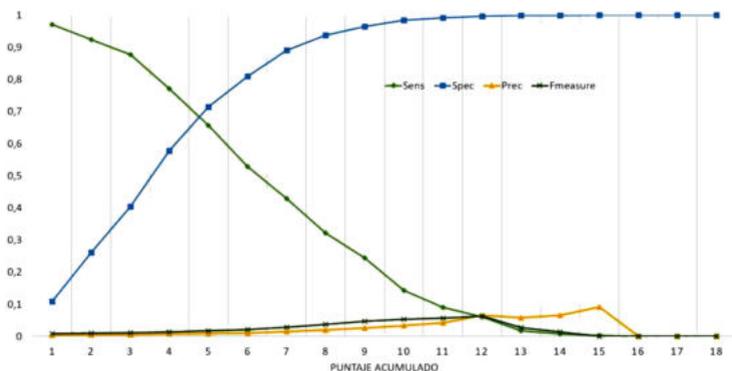


Figura 4. Métricas de evaluación según puntaje de corte.

métodos muy sensibles pueden sobre alertar al personal, perdiéndose el objetivo del modelo. Para afrontar este *trade-off* los modelos de puntuación poseen más de dos categorías de riesgo. Así, en las categorías de bajo riesgo se exige una alta sensibilidad y en las categorías de alto riesgo se espera una especificidad mayor. Al observar la Figura 4 es posible observar que los umbrales que cumplen esta condición son: entre 0 y 2, para el caso Sin Riesgo; 3 y 4 para Riesgo Bajo; 5 y 6 para Riesgo Medio; y sobre 7 para riesgo alto.

### DISCUSIÓN

El método propuesto en este trabajo asume que los registros de signos vitales deben tratarse como anormales si se encuentran en los extremos de las distribuciones en un grupo representativo de

pacientes hospitalizados. Esto permitió encontrar diferentes combinaciones de percentiles por signo vital que definen intervalos de riesgo. Diferentes combinaciones fueron estudiadas encontrándose una que presenta un mejor comportamiento respecto de su capacidad de predecir un traslado no programado a UCI en un intervalo de 8 horas. Esto se muestra por medio de los resultados obtenidos en Áreas bajo la curva ROC ( $AUC = 0,706$ ). Los umbrales de corte presentados cumplen con la propiedad de ser altamente sensibles en las categorías de riesgo menor (0,877 en los niveles de riesgo) y poseer mayor especificidad en las categorías de riesgo mayor (0,89 en la categoría de Alto Riesgo).

El SAT desarrollado en este trabajo tiene la característica de ser un sistema híbrido, que incluye

conocimiento experto (en la selección de las variables) y basado en datos para la determinación de los umbrales y rangos de normalidad.

Respecto a las limitaciones del modelo desarrollado, la data utilizada para la evaluación es desbalanceada y escasa. Esto se debe a que no se contó con más información para alimentar el modelo y que los eventos de alto riesgo no son tan frecuentes. La segunda limitación es que el estudio se basa en el análisis retrospectivo de pacientes hospitalizados. Otra limitación importante de este estudio es la suposición de que los valores de SAT calculados por paciente son independientes. Esta es una suposición habitual en los trabajos que evalúan SAT, pero esta suposición puede no ser válida en la práctica. Es decir, una medición de signos vitales en un instante puede estar altamente correlacionada con mediciones previas.

A pesar de las limitaciones discutidas anteriormente, los resultados del presente estudio son claros y a la vez de utilidad para el hospital. Nuestro enfoque supervisado y basado en datos permite un ajuste simple del modelo en la medida en que nuevos datos estén disponibles. Nuestro enfoque se puede ampliar fácilmente para incluir nuevas variables, como el apoyo ventilatorio y el tiempo de llene capilar, sin tener que volver a entrenar el modelo para los signos vitales descritos. El mecanismo propuesto es factible de ser aplicado a subpoblaciones, como por ejemplo pacientes con patologías específicas, si se dispone de la data necesaria. Además, el enfoque permite calcular tanto los percentiles como los umbrales para otras poblaciones, por ejemplo, pacientes adultos.

## CONCLUSIONES

Nuestro enfoque estadístico supervisado proporciona un método sencillo y reproducible para el rápido desarrollo y ajuste de los SAT. El SAT desarrollado permite identificar el deterioro clínico en pacientes pediátricos utilizando 8 variables: Frecuencia Cardíaca, Frecuencia respiratoria, Saturación de oxígeno, Temperatura corporal y Presión arterial Sistólica.

Todas las variables seleccionadas inicialmente aportan en la estimación del score de riesgo, por lo que ninguna de estas se descartó. Pese a que, además, se disponía de la PAD y PAM, incluir estas

variables no aporta en una mejora en los modelos estudiados, desde el punto de la capacidad predictiva del modelo, por lo que no se incluyeron en el grupo de variables seleccionadas inicialmente. Incluir más variables, puede representar una posible extensión de la presente investigación, pero aumentar el número de variables consideradas complicaría el registro de signos vitales desde el punto de vista de la capacitación de personal, su incorporación en los protocolos de calidad y seguridad del paciente, así como en la modificación del sistema de registro digital actualmente implementado en el hospital.

El modelo desarrollado tiene la característica de ser simple y rápido en lo que respecta al registro de signos. Además, la construcción y evaluación del modelo son entendibles y auditables, lo que lo vuelve un sistema adecuado para su uso en ambientes clínicos reales. Lo anterior representa una gran ventaja del SAT desarrollado en comparación con otras estrategias como herramientas de *Machine Learning* que, si bien ha mostrado mejores resultados desde el punto de vista de su capacidad predictiva, se caracterizan por ser modelos difíciles de auditar y difíciles de comprender por personal no especializado en estos modelos.

Esta herramienta puede otorgar una visión general del estado de los pacientes de cualquier unidad de hospitalización, con el fin de generar estrategias de mitigación del riesgo. Así, en pacientes con mayor riesgo detectado se recomienda incrementar la frecuencia de monitoreo o realizar intervenciones específicas por personal capacitado.

Los enfoques de riesgo de los pacientes en salud requieren de una mirada interdisciplinaria. La informatización de la salud y los avances tecnológicos recientes son una oportunidad para lograr desarrollar herramientas que faciliten la colaboración activa entre disciplinas como la medicina y la ingeniería. Los estadios de desarrollo más avanzados en ambas disciplinas se alcanzarán únicamente a través de investigaciones que requieran la interrelación de conocimientos y herramientas provenientes de diferentes miradas.

Se requiere analizar el comportamiento del modelo con data proveniente de diferentes centros de atención, como señala [1]. Además, utilizando este método es posible analizar un mayor número de

combinaciones de percentiles. Este análisis requiere de aumentar la capacidad de procesamiento, por lo que se considera como una extensión a este trabajo.

### AGRADECIMIENTOS

Los autores desean agradecer a Carlos Wolff, Dra. María Elisa Nalegach, Dra. María Begoña Yarza y a Eduardo Alcaina por su cooperación en esta investigación. Esta investigación es parcialmente financiada por CONICYT, Programa de Formación de Capital Humano avanzado (CONICYT-PCHA/Doctorado Nacional/2015-21150115).

### REFERENCIAS

- [1] N. Alam, E. Hobbelink, A. van Tienhoven, P. van de Ven, E. Jansma and P. Nanayakkara. "The impact of the use of the early warning score (EWS) on patient outcomes: A systematic review". *Resuscitation*. Vol. 85 N° 5, pp. 587-594. 2014.
- [2] M. Capan, J.S. Ivy, T. Rohleder, J. Hickman and J. M. Huddleston. "Individualizing and optimizing the use of early warning scores in acute medical care for deteriorating hospitalized patients". *Resuscitation*. Vol. 93, pp. 107-112. 2015.
- [3] A.R. de Caen, M.D. Berg, L. Chameides, C.K. Gooden, R.W. Hickey, H.F. Scott, R.M. Sutton, J.A. Tijssen, A. Topjian, E.W. van der Jagt, S.M. Schexnayder and R.A. Samson. "Part 12: Pediatric advanced life support: 2015 american heart association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care". *Circulation*. Vol. 132 N° 18 Suppl 2, pp. S526-S542. 2015.
- [4] C. Downey, W. Tahir, R. Randell, J. Brown and D. Jayne. "Strengths and limitations of early warning scores: A systematic review and narrative synthesis". *International Journal of Nursing Studies*. Vol. 76, pp. 106-119. 2017.
- [5] H. Duncan, J. Hutchison and C.S. Parshuram. "The pediatric early warning system score: A severity of illness score to predict urgent medical need in hospitalized children". *Journal of Critical Care*. Vol. 21 N° 3, pp. 271-278. 2006.
- [6] C.M. Freitas, E.P. Preto and C.A.F. Nascimento. "Nursing interventions for the early detection of ward patients' clinical deterioration: an integrative review". *Revista de Enfermagem*. Vol. 4 N° 14, pp. 121-130. 2017.
- [7] S. Gerry, J. Birks, T. Bonnici, P.J. Watkinson, S. Kirtley and G.S. Collins. "Early warning scores for detecting deterioration in adult hospital patients: A systematic review protocol". *BMJ Open*. Vol. 7 N° 12. 2017.
- [8] J. Kjellberg, R.T. Wolf, M. Kruse, S.R. Rasmussen, J. Vestergaard, K.J. Nielsen and K. Rasmussen. "Costs associated with adverse events among acute patients". *BMC Health Services Research*. Vol. 17 N° 1, pp. 651. 2017.
- [9] M.D. Le Lagadec and T. Dwyer. "Scoping review: The use of early warning systems for the identification of in-hospital patients at risk of deterioration". *Australian Critical Care*. Vol. 30 N° 4, pp. 211-218. 2017.
- [10] M.C. McLellan and J.A. Connor. "The cardiac children's hospital early warning score (c-chews)". *Journal of Pediatric Nursing*. Vol. 28 N° 2, pp. 171-178. 2013.
- [11] P. McQuillan, S. Pilkington, A. Allan, B. Taylor, A. Short, G. Morgan, M. Nielsen, D. Barrett and G. Smith. "Confidential inquiry into quality of care before admission to intensive care". *BMJ*. Vol. 316 N° 7148, pp. 1853-1858. 1998.
- [12] R. Morgan, F. Lloyd-Williams, M. Wright and R. Morgan-Warren. "An early warning scoring system for detecting developing critical illness". *Clin Intensive Care*. Vol. 8, pp. 100. 1997.
- [13] C.S. Parshuram, J. Hutchison and K. Middaugh. "Development and initial validation of the bedside paediatric early warning system score". *Critical Care*. Vol. 13 N° 4, pp. R135. 2009.
- [14] D.R. Prytherch, G.B. Smith, P.E. Schmidt and P.I. Featherstone. "Views-towards a national early warning score for detecting adult inpatient deterioration". *Resuscitation*. Vol. 81 N° 8, pp. 932-937. 2010.
- [15] A.G. Rapsang and D.C. Shyam. "Scoring systems in the intensive care unit: A compendium. Indian journal of critical care medicine: Peer-reviewed". *Indian Society of Critical Care Medicine*. Vol. 18 N° 4, pp. 220-228. 2014.
- [16] Royal College of Physicians. "National early warning score (NEWS) - standardising the

- assessment of acute-illness severity in the NHS”. Report of a working party. July 2012.
- [17] R.M. Schein, N. Hazday, M. Pena, B.H. Ruben and C.L. Sprung. “Clinical antecedents to in-hospital cardiopulmonary arrest”. *Chest*. Vol. 98 N° 6 pp. 1388-1392. 1990.
- [18] G.B. Smith, D.R. Prytherch, P. Meredith, P.E. Schmidt and P.I. Feath-erstone. “The ability of the national early warning score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated in-tensive care unit admission and death”. *Resuscitation*. Vol. 84 N° 4, pp. 465-470. 2013.
- [19] C.P. Subbe, M. Kruger, P. Rutherford and L. Gemmel. “Validation of a modified early warning score in medical admissions”. *QJM: Monthly Journal of the Association of Physicians*. Vol. 94 N° 10, pp. 521-526. 2001.
- [20] L. Tarassenko, D. A. Clifton, M.R. Pinsky, M.T. Hravnak, J.R. Woods and P.J. Watkinson. “Centile-based early warning scores de-rived from statistical distributions of vital signs”. *Resuscitation*. Vol. 82 N° 8, pp. 1013-1018. 2011.
- [21] P.J. Watkinson, M.A. Pimentel, D.A. Clifton and L. Tarassenko. “Manual centile-based early warning scores derived from statistical distributions of observational vital-sign data”. *Resuscitation*. Vol. 129, pp. 55-60. 2018.
- [22] P. Wolff, S.A. Ríos and M. Graña. “Setting up standards: A methodological proposal for pediatric triage machine learning model construction based on clinical outcomes”. *Expert Systems with Applications*. Vol. 138, pp. 112788. 2019.

## APENDICE A. DETALLE MÉTRICAS DE EVALUACIÓN PARA LA PRUEBA 6

El detalle de las métricas de evaluación para la Prueba 6 se muestra en la Tabla A.2.

Tabla 2. Resultados por puntaje de corte en la prueba 6.

|           | <b>Sens</b> | <b>Spec</b> | <b>Prec</b> | <b>F-score</b> |
|-----------|-------------|-------------|-------------|----------------|
| 0 Puntos  | 1,00        | 0,00        | 0,00        | 0,00           |
| 1 Punto   | 0,97        | 0,11        | 0,00        | 0,01           |
| 2 Puntos  | 0,92        | 0,26        | 0,00        | 0,01           |
| 3 Puntos  | 0,88        | 0,40        | 0,01        | 0,01           |
| 4 Puntos  | 0,77        | 0,58        | 0,01        | 0,01           |
| 5 Puntos  | 0,66        | 0,71        | 0,01        | 0,02           |
| 6 Puntos  | 0,53        | 0,81        | 0,01        | 0,02           |
| 7 Puntos  | 0,43        | 0,89        | 0,01        | 0,03           |
| 8 Puntos  | 0,32        | 0,94        | 0,02        | 0,04           |
| 9 Puntos  | 0,24        | 0,97        | 0,03        | 0,05           |
| 10 Puntos | 0,14        | 0,98        | 0,03        | 0,05           |
| 11 Puntos | 0,09        | 0,99        | 0,04        | 0,06           |
| 12 Puntos | 0,06        | 1,00        | 0,07        | 0,06           |
| 13 Puntos | 0,02        | 1,00        | 0,06        | 0,03           |
| 14 Puntos | 0,01        | 1,00        | 0,06        | 0,01           |
| 15 Puntos | 0,00        | 1,00        | 0,09        | 0,00           |
| 16 Puntos | 0,00        | 1,00        | 0,00        | 0,00           |
| 17 Puntos | 0,00        | 1,00        | 0,00        | 0,00           |
| 18 Puntos | 0,00        | 1,00        | 0,00        | 0,00           |